CUSTOMER NO.: 24737

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of A. KOOIMAN	)	Examiner: P. SHAH
Serial No.: 10/532,919	)	Art Unit: 2626
Filed: January 16, 2006	)	Confirmation: 4947
For: **METHOD FOR OPERATING A SPEECH RECOGNITION SYSTEM**	)	
Date of Examiner's Answer: February 24, 2010	)	
Attorney Docket No.: PHDE020239US / PKRX 200111US01	)	Cleveland, OH 44114 April 14, 2010

REPLY BRIEF

Commissioner For Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This Reply Brief is responsive to the Examiner's Answer of February 24, 2010. This Brief is responsive to new points raised in the Examiner's Answer.

EXAMINER'S ANSWER SECTIONS (1)-(8)

The Appellant and the Examiner appear to be in agreement concerning Sections (1)-(8) of the Appeal Brief and the Examiner's Answer.

EXAMINER'S ANSWER SECTION (9)

Section (9) of the Examiner's Answer merely repeats the Final Rejection, verbatim. Accordingly, Section (9) raises no new issues to which the applicant is entitled to respond.

EXAMINER'S ANSWER SECTION (10) – REPLY TO RESPONSE TO ARGUMENT

A. Claims 1, 4, and 10 are not anticipated by Polikaitis

Polikaitis discloses four thresholds which are applied in steps **230**, **240**, **250**, and **260**. Step **230** determines whether the user spoke over the start of the speech acquisition window, error 1 (column 6, lines 29-32). Step **240** determines whether the user spoke over the end of the speech acquisition window, error 2 (column 7, lines 15-18). Step **250** determines whether the user spoke too loudly, error 3 (column 8, lines 3-5). Step **260** determines whether the user spoke too softly, error 4 (column 8, lines 46-49). If the user makes any of these four user errors, Polikaitis communicates the error to the user via a communication interface circuitry.

By contrast, **claim 1** calls for determining a reception quality value or a noise value which represents the current reception quality. Reception quality, as used in the present application, relates the quality of the transmission channel, such as a telephone link, or a background noise level which is dependent on the environment in which the user is present (page 2, lines 10-13). Thus, claim 1 is concerned with the quality of the received signal; whereas, Polikaitis is concerned with user error.

The applicant must disagree with the Examiner's assertion that talking over the beginning of a window, talking past the end of the window, speaking too loudly, or speaking too softly constitute reception quality or a noise value which represents reception quality. The Examiner indicates, at page 21 of the Examiner's Answer, that claim 1 sets forth either switching the mode or outputting the alert signal, and that he is asserting that Polikaitis discloses the second option, as opposed to the first option.

Polikaitis does not output the alert signal when the reception quality value drops below a given reception quality threshold. Rather, Polikaitis instructs the user to change his/her behavior, particularly that the user should wait for the beginning of the speech window before talking, should not talk past the speech window, should talk softer, or should talk louder.

Accordingly, it is submitted that **claim 1** and **claims 2-8 and 10 dependent therefrom** are not anticipated by Polikaitis.

B. Claim 3 Distinguishes Patentably Over Polikaitis

Claim 2 further limits claim 1 by requiring switching between modes of operation. Specifically, when read in combination with its parent claim, claim 2 calls for switching the system to a mode of operation that is less-sensitive to noise, and resetting the speech recognition system to the previous (more noise sensitive) mode of operation when the reception quality exceeds the threshold or the noise level drops below the threshold. Thus, claim 2 requires switching between modes of operation.

Polikaitis does not switch modes of operation between a mode which is more sensitive to noise and a mode which is less sensitive to noise. Note the last paragraph of page 21 of the Examiner's Answer, in which the Examiner effectively concedes that Polikaitis does not switch modes. The Examiner's Answer on pages 23 and 24 does not identify a mode of Polikaitis which is more sensitive to noise and a mode which is less sensitive to noise.

On page 23, the Examiner asserts that instructing the user to repeat the voice instructions and to speak after the start of the speech acquisition window is a change of operation modes (column 6, lines 62-65). It is submitted that asking a user to repeat the input is not a change of modes, much less a change of modes between a more noise sensitive mode and a less noise sensitive mode. Similarly, prompting the user to repeat the speech instruction if he/she speaks over the end of a window is not a change of modes. Rather, asking the user to repeat a voice instruction in Polikaitis is clearly performed in the same mode for both the initial instruction and the repeated instruction. Nowhere does the Examiner point to any portion of Polikaitis which suggests changing modes, much less between more and less noise sensitive modes. Polikaitis leaves it to the user to speak more loudly, speak more softly, start speaking at the right time, or stop speaking at the right time, but does not change the hardware or software to operate in a different mode of operation. Accordingly, it is submitted that **claim 2** is not anticipated by Polikaitis.

C. Claim 3 Distinguishes Patentably Over Polikaitis, Nguyen, and Crane

Claim 3, when read with its parent claims, calls for the switching of modes, particularly switching from a barge-in mode to a non-barge-in mode when the reception quality drops below the threshold or the noise exceeds the threshold. By

way of background, the barge-in mode is a mode in which the user can barge-in on the speech output of the speech recognition system. More specifically, as explained in the third paragraph of page 4 of the present application, in the barge-in mode, the user does not have to wait for the prompt of the speech recognition system to finish before the user starts speaking. In the other mode, the user must wait for the end of the prompt. Polikaitis does not have a barge-in mode. Rather, the only mode described in Polikaitis requires that the user not start speaking until the speech acquisition window begins.

By contrast, Nguyen has a barge-in mode, and as the Examiner points out on page 25, determines when a barge-in has occurred. More specifically, as described at column 4, lines 46-53, Nguyen determines if a barge-in has occurred and, in response to such a barge-in, terminates the prompt and begins the recognition process.

Thus, Polikaitis has no barge-in mode. Nguyen has only a barge-in mode. Neither reference individually or in combination discloses or teaches that a system should have both modes, much less switch between these two modes, much less that switching should be a function of the reception quality value or noise value crossing a threshold.

Crane does not cure this shortcoming. Rather, Crane is concerned with determining when a barge-in has occurred. Note column 5, lines 22-43. That is, Crane, like Nguyen, stops the prompt in response to detecting a barge-in, i.e., operates in the barge-in mode at all times. Crane recognizes that all noises are not barge-ins, e.g., a cough or a door shutting should not be recognized as a barge-in. Crane is concerned with distinguishing these false barge-ins from true barge-ins. Thus, Crane, like Nguyen, functions only in a barge-in mode. Crane, like Nguyen, does not suggest two modes of operation that the system switches between, much less that one of the modes is a barge-in mode. Rather, Crane, Nguyen, and Polikaitis each only have a single mode of operation. The single mode of operation in Crane and Nguyen is the barge-in mode. The single mode in Polikaitis is not a barge-in mode. None of the references teach or put the reader in possession of the idea that the system should switch back and forth between modes, much less in and out of a barge-in mode.

Accordingly, it is submitted that **claim 3** distinguishes patentably over the references of record.

D. Claim 5 Distinguishes Patentably Over Polikaitis and Gerven

Gerven is concerned with techniques for determining whether speech is being received. More particularly, Gerven compares three algorithms for distinguishing speech from background noise. As the Examiner notes, section 2.3 of Gerven analyzes background noise to determine an adaptive threshold or baseline for distinguishing background noise from voice.

Claim 5 calls for the reception quality value or the noise value which is determined and used in the steps of claim 1 to be determined on the basis of a background signal prior to the beginning of an utterance. By contrast, Gerven listens substantially continuously to ascertain whether noise or speech is being received. Gerven does not teach determining a reception quality value or a noise value which is used to switch a mode of operation or to output an alert signal on the basis of the analysis of a background signal received prior to the beginning of an utterance. Rather, Gerven performs his analysis only to determine if an utterance has started.

Combining Polikaitis with Gerven would still not result in the limitations of claim 5. Gerven would only add to Polikaitis one or more techniques for determining when the user started speaking, which Polikaitis would use to determine whether the commencement of speaking was in the voice acquisition window or starting before it. Neither Gerven nor Polikaitis, individually or taken together, suggest or teach that this background noise prior to the beginning of an utterance should be used to switch between modes or to trigger an alarm. Indeed, in Polikaitis, none of error 1-error 4 are triggered based on a background signal before a user starts speaking. Rather, errors 1 and 2 are based on when the user starts or stops speaking and errors 3 and 4 are based on whether the user is speaking too loudly or too softly. Again, neither Polikaitis, nor Gerven, nor the combination thereof, teach using background noise as the basis for switching modes or generating an alert.

E. Claims 6 and 7 Are Patentable Over Polikaitis and Marx

Claim 6 is specific to the embodiment of Figure 2, in which the reception quality detector **6** and the comparator **11** are both in the voice activity detector **5**.

By contrast, Marx cited by the Examiner determines a confidence parameter indicative of how confident it is that the speech was interpreted properly.

The confidence value cannot be determined at a voice activity detector. Rather, the confidence level can only be determined after the speech is analyzed. Because Marx cannot determine the confidence level until the speech is analyzed, Marx not only does not teach, but cannot determine the confidence level with a voice activity detector.

Accordingly, it is submitted that **claim 6** and **claim 7 dependent therefrom** distinguish patentably over the references of record.

F. Claim 8 Distinguishes Patentably Over Polikaitis, Van Buskirk and Steinbrenner

The applicant and the Examiner appear to be in agreement that Steinbrenner relates to a telephony system or a network which can transmit diagnostic information to one of the telephones on the network, e.g., that the phone is off the hook. Steinbrenner is concerned with system malfunctions. Polikaitis is concerned with user errors, i.e., talking too soon, talking too long, talking too loud, or talking too soft. It is unclear how Steinbrenner would detect any of these situations. Indeed, it is submitted that there is no apparent relationship or teaching to combine Polikaitis and Steinbrenner. It is unclear why one would be motivated to combine them and what one would be hoping to achieve by combining them. Not only is there no teaching to combine these references, but the combination would not meet the limitations of **claim 8** because Polikaitis and Steinbrenner would be addressing different problems in different ways and interacting with different parts of a system to achieve different results.

Accordingly, it is submitted that claim 8 distinguishes patentably over the references of record.

G. Claims 9 and 11-13 Distinguish Patentably Over Polikaitis, Marx, and Bridges

As discussed above, Polikaitis determines whether the user starts speaking too soon, speaks too long, speaks too loudly, or speaks too softly. Marx generates a confidence signal based on a similarity of the translated answers to the expected answers. If one were to combine Marx with Polikaitis, it is submitted that one would add another step after the voice signal has been transcribed into language which determines whether or not the answer is an appropriate answer. If the answer

does not make sense, Marx, like Polikaitis, would ask the user to repeat it. Thus, the Marx process is downstream from the Polikaitis technique. Marx does not go to the reception quality value, but rather goes to the propriety of the answer. Thus, neither Polikaitis nor Marx determine a reception quality value or a noise value as called for by **claim 9**.

Bridges, like Gerven, determines if the incoming signal includes direct speech from the user, in which case the speech generator is deactivated and the speech recognizer is activated. That is, Bridges merely determines whether the user is talking to the system and if so, readies it to receive the input speech.

If one were to add the concepts of Bridges to Polikaitis, one would merely add the idea of determining whether a received signal is direct speech from the user or not, and if it is direct speech from the user, activates the speech recognizer and deactivates the speech generator. This would not cure any of the previously-discussed shortcomings of Polikaitis or Marx. It would just prevent the system from reacting to noise other than the user's voice.

For these reasons and the reasons set forth in the Appeal Brief, it is submitted that Polikaitis, Marx, and Bridges do not put the reader in possession of the speech recognition system defined by **claim 9** or **claims 11, 12, and 13 dependent therefrom**.

H. Claim 12 Distinguishes Patentably Over Polikaitis, Marx, and Bridges

Claim 9 calls for a control means adapted to switch the speech recognition system over to a mode of operation which is less sensitive to noise or output alert signal to the user. **Claim 12** further clarifies that the control means includes a barge-in switching unit. A barge-in switching unit is not going to output an alert signal to a user. Indeed, in Marx, detecting user speech does not cause an alert signal to be output. Rather, the detection of user speech disables the speech synthesizer and enables the voice recognition software.

The claimed barge-in switching unit switches the mode of operation between a more noise sensitive and less noise sensitive mode of operation and to perform such switching based on a reception quality value or a noise value. Marx does not disclose more and less noise sensitive modes of operation, much less switching between them, much less using a barge-in switching unit.

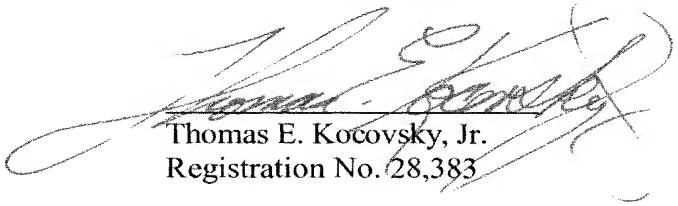
Because the Examiner not having discussed Bridges in the Examiner's Answer relative to claim 12, no rebuttal regarding Bridges is deemed appropriate.

CONCLUSION

For the reasons set forth above, in the Appeal Brief and in the additional rebuttal set forth above, it is submitted that claims 1-13 are not anticipated by and distinguish patentably over the references of record.

An early Decision reversing the Examiner's rejections of all claims is requested.

Respectfully submitted,



Thomas E. Kocovsky, Jr.
Registration No. 28,383

FAY SHARPE LLP
The Halle Building, 5th Floor
1228 Euclid Avenue
Cleveland, OH 44115-1843
Telephone: 216.363.9000 (main)
Telephone: 216.363.9122 (direct)
Facsimile: 216.363.9001
E-Mail: tkocovsky@faysharpe.com

Direct All Correspondence to:
Yan Glickberg, Reg. No. 51,742
US PHILIPS CORPORATION
P.O. Box 3001
Briarcliff Manor, NY 10510-8001
(440) 483-3455 (tel)
(440) 483-2452 (fax)